GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: December 13, 2002, 02:54:10 ; Search time 297 Seconds
(without alignments)
15218.049 Million cell updates/sec

Title: US-09-716-536-7
Perfect score: 2007
Sequence: 1 gtgcggtggagcgaaatttg.....aaaaaaaaaaaaaaaaaaa 2007

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2007	100.0	2007	24	AAD31618	Human TNF receptor
	2	1890.8	94.2	2065	20	AAX86754	cDNA 091-21A31 enc
	3	1887.6	94.1	2065	19	AAV29062	BRCA1 modulator pr
	4	1081	53.9	1975	24	AAD31619	Mouse TNF receptor
	5	892.8	44.5	1781	24	ABL39714	Human NS cDNA sequ
	6	176.4	8.8	7542	24	ABA95683	Human protein kina
c	7	148	7.4	148	22	AAS27719	DNA encoding novel
	8	60	3.0	60	24	ABN41009	Human spliced tran
	9	52	2.6	3489	21	AAA30290	Kaposi's sarcoma-a
	10	52	2.6	3489	22	AAF82901	Nucleotide sequenc
	11	52	2.6	3489	24	ABA93487	Kaposi's sarcoma-a
c	12	52	2.6	32207	20	AAV73805	KSHV LUR DNA (nucl
c	13	52	2.6	137507	19	AAV19941	KSHV long unique c
	14	51.8	2.6	4246	22	AAS60947	Human cancer agent
	15	51.8	2.6	7596	24	ABL66905	Lung cancer relate
	16	51.6	2.6	2108	24	ABL67774	Oesophagus cancer
c	17	50	2.5	50	22	AAL30971	Human SNP oligonuc
	18	48.4	2.4	7453	22	AAI58369	Human polynucleoti
	19	48.4	2.4	7501	22	AAI58370	Human polynucleoti
	20	48.4	2.4	7741	22	AAI60155	Human polynucleoti
	21	48.4	2.4	7741	22	AAI60156	Human polynucleoti
	22	47.8	2.4	5154	23	AAS84859	DNA encoding novel
	23	47.2	2.4	2004	18	AAT85356	Nephila clavipes s
c	24	46.6	2.3	475	22	ABA58819	Human foetal liver
c	25	46.6	2.3	475	22	ABA27737	Probe #6203 for ge
c	26	46.6	2.3	475	22	AAK06973	Human brain expres
c	27	46.6	2.3	475	22	AAK32709	Human bone marrow
c	28	46.6	2.3	475	22	AAI38524	Probe #7210 used t
c	29	46.6	2.3	475	24	ABS07505	Human genome-deriv
c	30	46.6	2.3	511	22	ABA71159	Human foetal liver
c	31	46.6	2.3	511	22	ABA37497	Probe #15963 for g
c	32	46.6	2.3	511	22	AAK19455	Human brain expres
c	33	46.6	2.3	511	22	AAK45444	Human bone marrow
c	34	46.6	2.3	511	22	AAI51389	Probe #20075 used
c	35	46.6	2.3	511	24	ABS19709	Human genome-deriv
	36	46.6	2.3	3809	24	ABK95303	Human prostate spe
	37	45.8	2.3	4181	22	AAD06778	Human haematopoiet
	38	45.8	2.3	4801	22	AAD06781	Human haematopoiet
	39	45.2	2.3	1820	22	AAF58611	Human RECAP polynu
c	40	45	2.2	1824	23	AAS81488	DNA encoding novel
	41	45	2.2	2850	23	AAS79695	DNA encoding novel
c	42	44.4	2.2	267	22	AAK19599	Human brain expres
c	43	44.4	2.2	267	22	AAK45604	Human bone marrow
c	44	44.4	2.2	267	24	ABS19876	Human genome-deriv
	45	44.2	2.2	693	23	AAS74240	DNA encoding novel

ALIGNMENTS

RESULT 1
AAD31618
ID AAD31618 standard; cDNA; 2007 BP.
XX
AC AAD31618;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human TNF receptor associated factor interacting protein (TRIP) cDNA.
XX
KW Human; tumour necrosis factor; TNF; TNF receptor associated factor; TRAF;
KW TRAF interacting protein; TRIP; cell activation; cell proliferation;
KW cell death; therapy; cytostatic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 103..1512
FT /*tag= a
FT /product= "Human TRIP"
XX
PN US6346605-B1.
XX
PD 12-FEB-2002.
XX
PF 31-MAR-1998; 98US-0052089.
XX
PR 01-APR-1997; 97US-042293P.
PR 07-APR-1997; 97US-042747P.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX

PI Lee SY, Choi Y;
XX
DR WPI; 2002-225005/28.
DR P-PSDB; AAE19853.
XX
PT New tumor necrosis factor receptor associated factor interacting
PT protein, useful for inhibiting NF-kappa B activation, and for
PT modulating signals responsible for cell activation, cell proliferation
PT and cell death -
XX
PS Example 2; Fig 8A; 37pp; English.
XX
CC The present invention relates to a tumour necrosis factor (TNF) receptor
CC associated factor (TRAF) interacting protein (TRIP), which is a regulator
CC capable of binding to TRAF2. TRIP is useful for inhibiting NF-kappa B
CC activation and for modulating signals responsible for cell activation,
CC cell proliferation and cell death. Thus, TRIP is useful for treating
CC diseases associated with altered cell proliferation and cell death. The
CC present sequence is human TRIP cDNA.
XX
SQ Sequence 2007 BP; 517 A; 518 C; 558 G; 414 T; 0 other;

Query Match 100.0%; Score 2007; DB 24; Length 2007;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GTGCGGTGGAGCGAAATTTGAAGCAAGCGGAGGCGGGGCGCTCTACGAAGCCGGACCTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTGCGGTGGAGCGAAATTTGAAGCAAGCGGAGGCGGGGCGCTCTACGAAGCCGGACCTGT 60

Qy      61 AGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCCTATCCGTGCTCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCCTATCCGTGCTCTG 120

Qy     121 TGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGCCATCCACTGCGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGCCATCCACTGCGGC 180

Qy     181 CACACCTTCCACTTGCAGTGCCTAATTCAGTCCTTTGAGACAGCACCAAGTCGGACCTGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CACACCTTCCACTTGCAGTGCCTAATTCAGTCCTTTGAGACAGCACCAAGTCGGACCTGC 240

Qy     241 CCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATAAGCTCTTCTTTGATCTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATAAGCTCTTCTTTGATCTT 300

Qy     301 GCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGAATGAACTGGACAATGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGAATGAACTGGACAATGTC 360

Qy     361 AGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCCAGGTCATCATCGACACT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCCAGGTCATCATCGACACT 420

Qy     421 CTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTCTGCAGCAGGCCTTGGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTCTGCAGCAGGCCTTGGGC 480

Qy     481 AAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAG 540

Qy     541 GATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATG 600

Qy     601 GAGCAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAGCAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGAC 660

Qy     661 ATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAA 720

Qy     721 GAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGG 780

Qy     781 AAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACTCTGAATTGGATCAGGCC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACTCTGAATTGGATCAGGCC 840

Qy     841 AAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTGACAAGGAAATCATGAGC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTGACAAGGAAATCATGAGC 900

Qy     901 CTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGCCACCAGTGGCCAGTGAG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGCCACCAGTGGCCAGTGAG 960

Qy     961 ACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGGTGAATCTGAAGCTCCGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGGTGAATCTGAAGCTCCGC 1020

Qy    1021 CGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTGATGTGGATACTCCCCCA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTGATGTGGATACTCCCCCA 1080

Qy    1081 GCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTTGCCTAGAGAAGTCACAC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTTGCCTAGAGAAGTCACAC 1140

Qy    1141 TCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCAGGAAGGAGTCCCAGCTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCAGGAAGGAGTCCCAGCTC 1200

Qy    1201 TCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAACTGGTTGGTGCCTTCCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAACTGGTTGGTGCCTTCCCT 1260

Qy    1261 ATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAAAGGCCCAGGTCAGAGTCC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAAAGGCCCAGGTCAGAGTCC 1320

Qy    1321 TCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCGGTGGCCGGACAAAATTC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCGGTGGCCGGACAAAATTC 1380

Qy    1381 ATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTAAGCCCAAGACCAAGGTT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 ATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTAAGCCCAAGACCAAGGTT 1440

Qy    1441 AAGCAGAGGGTGAGGGTGAAGACCGTGCCTTCTCTCTTCCAGGCCAAGCTGGACACCTTC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AAGCAGAGGGTGAGGGTGAAGACCGTGCCTTCTCTCTTCCAGGCCAAGCTGGACACCTTC 1500

Qy    1501 CTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACATGCCTGCAACTTGTAGG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACATGCCTGCAACTTGTAGG 1560

Qy    1561 TCAAGGACTGTCCAGGCAGGGTTTGTGGACAGAGCCCTACTTTCGGGACCAGCCTGAGGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TCAAGGACTGTCCAGGCAGGGTTTGTGGACAGAGCCCTACTTTCGGGACCAGCCTGAGGT 1620

Qy    1621 GTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAGAGACTGCTCTTCCTGCCCT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAGAGACTGCTCTTCCTGCCCT 1680

Qy    1681 CACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAGCCCACTGATCCTGTCAGCA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 CACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAGCCCACTGATCCTGTCAGCA 1740

Qy    1741 GGTCCTGCTCTGTTGCCAGGCTCTTGTTTATAGCCATGATCAGATGTGGTCAGACTCTTT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GGTCCTGCTCTGTTGCCAGGCTCTTGTTTATAGCCATGATCAGATGTGGTCAGACTCTTT 1800
```

```
Qy   1801 CTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGTGGACCAGAGTGCTTGAGGCA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGTGGACCAGAGTGCTTGAGGCA 1860

Qy   1861 TCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGACTTGCTTCTAGCATAGCCTGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGACTTGCTTCTAGCATAGCCTGG 1920

Qy   1921 GCCAAGCAGGGTGGGGAATGGAGGATAGACATGGGATGTATGGAGAGGATGGAAGATTTT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GCCAAGCAGGGTGGGGAATGGAGGATAGACATGGGATGTATGGAGAGGATGGAAGATTTT 1980

Qy   1981 CCCGAAAAAAAAAAAAAAAAAAAAAAAA 2007
          ||||||||||||||||||||||||||||
Db   1981 CCCGAAAAAAAAAAAAAAAAAAAAAAAA 2007
```

```
RESULT 2
AAX86754
ID   AAX86754 standard; cDNA; 2065 BP.
XX
AC   AAX86754;
XX
DT   27-OCT-1999  (first entry)
XX
DE   cDNA 091-21A31 encoding a BRCA1 modulator protein.
XX
KW   Modulator protein; BRCA1; tumour suppressor protein; breast cancer;
KW   ovarian cancer; cell growth; cell proliferation; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             103..1512
FT                   /*tag=  a
XX
PN   US5948643-A.
XX
PD   07-SEP-1999.
XX
PF   13-AUG-1997;   97US-0968751.
XX
PR   13-AUG-1997;   97US-0968751.
XX
PA   (ONYX-) ONYX PHARM INC.
XX
PI   Lingenfelter C,  Polakis PG,  Rubinfeld B,  Vuong TT;
XX
DR   WPI; 1999-517952/43.
DR   P-PSDB; AAY30149.
XX
PT   Modulator proteins that bind to and modulate the activity of the
PT   BRCA1 tumour suppressor gene product, useful for the treatment of
PT   ovarian and breast cancer
XX
PS   Claim 1; Fig 1; 35pp; English.
XX
CC   The present sequence encodes a modulator protein, that binds to and
CC   modulate the activity of the BRCA1 gene product (BRCA1). The BRCA1
CC   protein has been characterized as a tumour suppressor protein.
CC   Alterations in the amino acid sequence of BRCA1 causes breast and ovarian
CC   cancers by removing the controls on cell growth and proliferation.
CC   Research has shown that different regions on the BRCA1 molecule have
CC   different effects on cell growth and tumour suppression (e.g. full length
CC   truncated BRCA1 has no effect on breast cancer cell growth but will
CC   inhibit ovarian cancer cell growth). It has been suggested that different
CC   host cell factors (e.g. proteins) interact with different regions of the
CC   BRCA1 to control its function. The identification of these proteins
CC   (e.g. BRCA1MP) will facilitate the development of novel diagnostic
CC   methods and new therapeutics for identifying and treating cancers caused
CC   by changes in the expression or activity of BRCA1.
XX
SQ   Sequence 2065 BP; 561 A; 526 C; 561 G; 417 T; 0 other;
```

```
  Query Match             94.2%;  Score 1890.8;  DB 20;  Length 2065;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 1946;  Conservative    0;  Mismatches   17;  Indels    5;  Gaps    4;

Qy     44 TACGAAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCA 103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44 TACGAAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCA 103

Qy    104 TGCCTATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGG 163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 TGCCTATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGG 163

Qy    164 CCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCAGTCCTTTGAGACAG 223
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db    164 CCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCAGTGGTTTGAGACAG 223

Qy    224 CACCAAGTCGGACCTGCCCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATA 283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 CACCAAGTCGGACCTGCCCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATA 283

Qy    284 AGCTCTTCTTTGATCTTGCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGA 343
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db    284 AGCTCTTCTTTGATCTTGCCCAGGAGGAGGAGAATGTCTTGGATGCAGAATTCTTAAAGA 343

Qy    344 ATGAACTGGACAATGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCC 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 ATGAACTGGACAATGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCC 403

Qy    404 AGGTCATCATCGACACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTC 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 AGGTCATCATCGACACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTC 463

Qy    464 TGCAGCAGGCCTTGGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGT 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 TGCAGCAGGCCTTGGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGT 523

Qy    524 ACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGA 583
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    524 ACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAGGAGGCCCGCCGGCTCAGGA 583

Qy    584 GCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGG 643
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    584 GCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAGAGCCAGCGCCCTGAGGTGG 643

Qy    644 AGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACT 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 AGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACT 703

Qy    704 GTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGG 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 GTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGG 763

Qy    764 TGGCTGACAAGCTGAGGAAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACT 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TGGCTGACAAGCTGAGGAAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACT 823

Qy    824 CTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTG 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824 CTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTG 883

Qy    884 ACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGC 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 ACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGC 943

Qy    944 CACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGG 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 CACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGG 1003

Qy   1004 TGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063
```

```
Qy   1064 ATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTT 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 ATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTT 1123

Qy   1124 GCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 GCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCA 1183

Qy   1184 GGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAAC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 GGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAAC 1243

Qy   1244 TGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAAA 1303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1244 TGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAGA 1303

Qy   1304 GGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCG 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 GGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCG 1363

Qy   1364 GTGGCCGGACAAAATTCATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTA 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 GTGGCCGGACAAAATTCATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTA 1423

Qy   1424 AGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACCGTGCCTTCTCTCTTCCAGG 1483
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1424 AGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACAGTGCCTTCTCTCTTCCAGG 1483

Qy   1484 CCAAGCTGGACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACA 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 CCAAGCTGGACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACA 1543

Qy   1544 TGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGG--TTTGTGGACAGAGCCCTACT 1601
          |||||||||||||||||||||||||||||||||||||  |||||||||||||||| |||
Db   1544 TGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGGGTTTTGTGGACAGAGCCCCACT 1603

Qy   1602 TTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAG 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 TTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAG 1663

Qy   1662 AGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAG 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1664 AGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAG 1723

Qy   1722 CCCACTGATCCTGTCAGCAGGTCCTGCT-CTGTTGCCAGGCTCTTGTTTATAGCCATGAT 1780
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db   1724 CCCACTGATCCTGTCAGCAGGTCCTGCTCCTGTTGCCAGGCTCCTGTTTATAGCCATGAT 1783

Qy   1781 CAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGT 1840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1784 CAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGT 1843

Qy   1841 GGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGAC 1900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1844 GGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGAC 1903

Qy   1901 TTGCTTCTA-GCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGATAGACATGGGATGT 1959
          ||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1904 TTGCTTCTAGGCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGATAG-CATGGGATGT 1962

Qy   1960 ATGGAGAGGATGGAAGATTTTCCCGAAAAAAAAAAAAAAAAAAAAAA 2007
          ||||||||||||||||||||||  | |||| ||||  ||||||||||
Db   1963 ATGGAGAGGATGGAAGATTTTCATGTAAAATAAAATTAAAAAAAAAA 2010


RESULT 3
AAV29062
ID   AAV29062 standard; cDNA; 2065 BP.
XX
AC   AAV29062;
XX
DT   28-AUG-1998  (first entry)
XX
DE   BRCA1 modulator protein 091-21A31 cDNA.
XX
KW   BRCA1 modulator protein; 091-21A31; breast cancer antigen 1;
KW   tumour suppressor protein; diagnosis; therapy; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             103..1512
FT                   /*tag=  a
XX
PN   WO9810066-A1.
XX
PD   12-MAR-1998.
XX
PF   06-AUG-1997;   97WO-US13944.
XX
PR   04-SEP-1996;   96US-0025601.
XX
PA   (ONYX-) ONYX PHARM INC.
XX
PI   Ligenfelter C,  Polakis P,  Rubinfeld B,  Vuong TT;
XX
DR   WPI; 1998-193616/17.
DR   P-PSDB; AAW37881.
XX
PT   Breast cancer antigen 1 modulator protein - useful for diagnosing
PT   diseases involving unwanted cell growth, e.g. breast cancer, and for
PT   producing therapeutics for treatment of such diseases
XX
PS   Claim 5; Fig 1; 73pp; English.
XX
CC   This cDNA clone, designated 091-21A31 (ATCC 98141), codes for
CC   a 53 kDa BRCA1 modulator protein (see AAW37881) that binds to the
CC   tumour suppressor gene product BRCA1, and which is characterised by
CC   a zinc finger domain and a leucine zipper motif.  3 cDNA clones
CC   (see also AAV29063 and AAV29064) coding for BRCA1 modulator proteins
CC   (see AAW37881-83) were isolated from a HeLa cDNA library using a
CC   yeast two-hybrid assay with a GAL4-BRCA1(8-1293) fusion as bait.
CC   Vectors and host cells comprising the isolated nucleic acid
CC   sequences are claimed, as well as a process for producing BRCA1
CC   modulator protein by culturing these host cells.  BRCA1 modulator
CC   proteins and nucleic acids can be used to diagnose diseases
CC   involving unwanted cell growth, e.g. breast cancer, and to identify
CC   compounds that alter BRCA1 interaction with BRCA1 modulators for
CC   the treatment of such diseases.
XX
SQ   Sequence 2065 BP; 561 A; 528 C; 559 G; 417 T; 0 other;

  Query Match             94.1%;  Score 1887.6;  DB 19;  Length 2065;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1944;  Conservative    0;  Mismatches   19;  Indels    5;  Gaps    4;

Qy     44 TACGAAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCA 103
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     44 TACGAAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGACTCCAGCCATCA 103

Qy    104 TGCCTATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGG 163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 TGCCTATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGG 163

Qy    164 CCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCAGTCCTTTGAGACAG 223
          |||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db    164 CCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCAGTGGTTTGAGACAG 223

Qy    224 CACCAAGTCGGACCTGCCCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATA 283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 CACCAAGTCGGACCTGCCCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATA 283

Qy    284 AGCTCTTCTTTGATCTTGCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGA 343
          |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db    284 AGCTCTTCTTTGATCTTGCCCAGGAGGAGGAGAATGTCTTGGATGCAGAATTCTTAAAGA 343
```

```
Qy    344 ATGAACTGGACAATGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCC 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 ATGAACTGGACAATGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCC 403

Qy    404 AGGTCATCATCGACACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTC 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 AGGTCATCATCGACACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTC 463

Qy    464 TGCAGCAGGCCTTGGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGT 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 TGCAGCAGGCCTTGGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGT 523

Qy    524 ACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGA 583
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    524 ACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAGGAGGCCCGCCGGCTCAGGA 583

Qy    584 GCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGG 643
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    584 GCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAGAGCCAGCGCCCTGAGGTGG 643

Qy    644 AGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACT 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    644 AGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACT 703

Qy    704 GTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGG 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    704 GTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGG 763

Qy    764 TGGCTGACAAGCTGAGGAAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACT 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TGGCTGACAAGCTGAGGAAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACT 823

Qy    824 CTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTG 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    824 CTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTG 883

Qy    884 ACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGC 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 ACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGC 943

Qy    944 CACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGG 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 CACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGAGG 1003

Qy   1004 TGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063

Qy   1064 ATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTT 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 ATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTT 1123

Qy   1124 GCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 GCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCA 1183

Qy   1184 GGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAAC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 GGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAAC 1243

Qy   1244 TGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAAA 1303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   1244 TGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAGA 1303

Qy   1304 GGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCG 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 GGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCG 1363

Qy   1364 GTGGCCGGACAAAATTCATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTA 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 GTGGCCGGACAAAATTCATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTA 1423

Qy   1424 AGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACCGTGCCTTCTCTCTTCCAGG 1483
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1424 AGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACAGTGCCTTCTCTCTTCCAGG 1483

Qy   1484 CCAAGCTGGACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACA 1543
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 CCAAGCTGCACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACA 1543

Qy   1544 TGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGG--TTTGTGGACAGAGCCCTACT 1601
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||| |||||
Db   1544 TGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGGGTTTTGTGGACAGAGCCCCACT 1603

Qy   1602 TTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAG 1661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 TTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAG 1663

Qy   1662 AGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAG 1721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1664 AGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAG 1723

Qy   1722 CCCACTGATCCTGTCAGCAGGTCCTGCT-CTGTTGCCAGGCTCTTGTTTATAGCCATGAT 1780
          |||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db   1724 CCCACTGATCCTGTCAGCAGGTCCTGCTCCTGTTGCCAGGCTCCTGTTTATAGCCATGAT 1783

Qy   1781 CAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGT 1840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1784 CAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGT 1843

Qy   1841 GGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGAC 1900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1844 GGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGAC 1903

Qy   1901 TTGCTTCTA-GCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGATAGACATGGGATGT 1959
          |||||||||  |||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1904 TTGCTTCTAGGCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGATAG-CATGGGATGT 1962

Qy   1960 ATGGAGAGGATGGAAGATTTTCCCGAAAAAAAAAAAAAAAAAAAAAA 2007
          ||||||||||||||||||||||  | |||| |||| ||||||||||
Db   1963 ATGGAGAGGATGGAAGATTTTCATGTAAAATAAAATTAAAAAAAAAAA 2010
```

```
RESULT 4
AAD31619
ID   AAD31619 standard; cDNA; 1975 BP.
XX
AC   AAD31619;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Mouse TNF receptor associated factor interacting protein (TRIP) cDNA.
XX
KW   Mouse; tumour necrosis factor; TNF; TNF receptor associated factor; TRAF;
KW   TRAF interacting protein; TRIP; cell activation; cell proliferation;
KW   cell death; therapy; cytostatic; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             112..1524
FT                   /*tag= a
FT                   /product= "Mouse TRIP"
FT                   /transl_except= (pos:922..923, aa:Ser)
FT                   /note= "There is apparent deletion of one base
FT                   which alters the reading frame"
XX
PN   US6346605-B1.
XX
PD   12-FEB-2002.
XX
PF   31-MAR-1998;   98US-0052089.
XX
PR   01-APR-1997;   97US-042293P.
PR   07-APR-1997;   97US-042747P.
XX
```

PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Lee SY, Choi Y;
XX
DR WPI; 2002-225005/28.
DR P-PSDB; AAE19854.
XX
PT New tumor necrosis factor receptor associated factor interacting
PT protein, useful for inhibiting NF-kappa B activation, and for
PT modulating signals responsible for cell activation, cell proliferation
PT and cell death -
XX
PS Example 2; Fig 8B; 37pp; English.
XX
CC The present invention relates to a tumour necrosis factor (TNF) receptor
CC associated factor (TRAF) interacting protein (TRIP), which is a regulator
CC capable of binding to TRAF2. TRIP is useful for inhibiting NF-kappa B
CC activation and for modulating signals responsible for cell activation,
CC cell proliferation and cell death. Thus, TRIP is useful for treating
CC diseases associated with altered cell proliferation and cell death. The
CC present sequence is mouse TRIP cDNA.
CC Note: This sequence SEQ ID NO:8 is stated to be similar to the sequence
CC shown in the sequence listing. However the sequence shown in sequence
CC listing lacks few bases at the end of each line.
XX
SQ Sequence 1975 BP; 530 A; 488 C; 533 G; 424 T; 0 other;

Query Match 53.9%; Score 1081; DB 24; Length 1975;
Best Local Similarity 74.6%; Pred. No. 7.6e-283;
Matches 1504; Conservative 0; Mismatches 450; Indels 63; Gaps 9;

```
Qy       1 GTGCGGTGGAGCGAAATTTGAAGCAAGCGGAGGCGGGGC--GCTCTACGAAGCCGGACCT 58
Db       9 GTGCGGTGGAGCGAAATTTGAAGGAACCGGAGCGGTGGCCGGTTCCACCAAACTGTGTCT 68

Qy      59 GTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCCTATCCGTGCTC 118
Db      69 GTCTCTG-GCAGCTGGTTCCCTGGGCTGCTTGAGTCGAGCCATCATGCCTATCCTCTCTC 127

Qy     119 TGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGCCATCCACTGCG 178
Db     128 TGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGTGACGTGGCTGCCATCCACTGTG 187

Qy     179 GCCACACCTTCCACTTGCAGTGCCTAATTCAGTCCTTTGAGACAGCACCAAGTCGGACCT 238
Db     188 GCCACACTTTTCATCTGCAATGCCTAATCCAGTGGTTTGAGACAGCACCAAGTCGGACCT 247

Qy     239 GCCCACAGTGCCGAATCCAGGTTGGCAAAAGAACCATTATCAATAAGCTCTTCTTTGATC 298
Db     248 GCCCACAGTGTAGAATCCAGGTTGGCAAAAAGACTATTATAAACAAACTTTTCTTTGACC 307

Qy     299 TTGCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGAATGAACTGGACAATG 358
Db     308 TCGCCCAGGAAGAGGAGAATGTCTTGGATGCAGAATTCTTAAAGAATGAACTGGACAGCG 367

Qy     359 TCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCCAGGTCATCATCGACA 418
Db     368 TCAAAGCTCAGCTTTCCCAGAAAGACAGGGAGAAACGGGACAGCCAGGCCATTATCGACA 427

Qy     419 CTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTCTGCAGCAGGCCTTGG 478
Db     428 CTCTACGGGACACCCTGGAAGAACGCAATGCTACCGTGGAGTCCCTACAGAACGCCTTAA 487

Qy     479 GCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGC 538
Db     488 ACAAGGCAGAGATGCTGTGTTCCACCCTGAAAAAACAGATGAAGTTCCTGGAGCAGCGGC 547

Qy     539 AGGATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCA 598
Db     548 AGGATGAGACCAAACAAGCTCGGGAGGAGGCCCACCGACTCAAGTGCAAGATGAAAACCA 607

Qy     599 TGGAGCAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAG 658
```

```
Db     608 TGGAGCAAATTGAGCTCCTACTCCAGAGCCAGCGTTCTGAGGTGGAGGAGATGATTCGAG 667

Qy     659 ACATGGGTGTGGGACAGTCAGCGGTGGAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGA 718
Db     668 ACATGGGTGTGGGACAGTCAGCGGTGGAGCAGCTGGCTGTGTACTGCGTGTCCCTCAAGA 727

Qy     719 AAGAGTACGAGAATCTAAAAGAGGCACGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGA 778
Db     728 AAGAGTATGAGAATCTGAAGGAAGCTCGGAAGGCCACAGGGGAACTGGCTGACAGGTTGA 787

Qy     779 GGAAGGATTTGTTTTCCTCCAGAAGCAAGTTGCAGACAGTCTACTCTGAATTGGATCAGG 838
Db     788 AGAAGGATTTGGTGTCCTCTAGGAGCAAGTTGAAGACTCTCAACACTGAGCTGGATCAGG 847

Qy     839 CCAAGTTAGAACTGAAGTCAGCCCAGAAGGACTTACAGAGTGCTGACAAGGAAATCATGA 898
Db     848 CCAAGTTAGAACTGAGGTCAGCCCAGAAGGACTTACAAAGTGCTGACCAGGAGATCACGA 907

Qy     899 GCCTGAAAAAGAAGCTAACGATGCTGCAGGAAACCTTGAACCTGCCACCAGTGGCCAGTG 958
Db     908 GCCTAAGAAAGAAGCTGATGATCCTCCAGGGAACCTTGAGCCTGCCTCCGCGTACCAATG 967

Qy     959 AGACTGTCGACCGCCTGGTTTTAGAGAGCCCAGCCCCTGTGGA---GGTGAATCTGAAGC 1015
Db     968 AGACGGTCAGCCGCCTGGTTTTTGAGAGCCCAGCCCCTGTGGAGATGATGAACCCGAGGC 1027

Qy    1016 TCCGCCGGCCATCCTTCCGTGATGATATTGATCTCAATGCTACCTTTGATGTGGATACTC 1075
Db    1028 TTCACCAGCCACCCTTCGGTGATGAGATTGATCTCAATACCACCTTTGATGTAAATACCC 1087

Qy    1076 CCCCAGCCCGGCCCTCCAGCTCCCAGCATGGTTACTACGAAAAACTTTGCCTAGAGAAGT 1135
Db    1088 CTCCAACCCAGACCTCTGGCTCCCAGCATTGCCTCCCCAAGAAGCTGTGCCTGGAGAGGG 1147

Qy    1136 CACACTCCCCAATTCAGGATGTCCCCAAGAAGATATGCAAAGGCCCCAGGAAGGAGTCCC 1195
Db    1148 CACGCTCTCCCATGCAGAATGTCCTCAAGAAGGTGCACAAAGTCTCCAAGCCGGAGTCCC 1207

Qy    1196 AGCTCTCACTGGGTGGCCAGAGCTGTGCAGGAGAGCCAGATGAGGAACTGGTTGGTGCCT 1255
Db    1208 AGCTCTCACTGGGTGGCCAGCGATGTGTAGGAGAGCTAGATGAGGAACTGGCTGGTGCCT 1267

Qy    1256 TCCCTATTTTTGTCCGGAATGCCATCCTAGGCCAGAAACAGCCCAAAAGGCCCAGGTCAG 1315
Db    1268 TCCCTCTCTTCATCCGGAATGCTGTCCTGGGTCAGAAACAGCCCAACAGGACCACAGCAG 1327

Qy    1316 AGTCCTCTTGCAGCAAAGATGTGGTAAGGACAGGCTTCGATGGGCTCGGTGGCCGGACAA 1375
Db    1328 AATCCCGAAGCAGCACAGATGTGGTAAGAATAGGCTTTGATGGGCTTGGAGGACGAACAA 1387

Qy    1376 AATTCATCCAGCCTACTGACACAGTCATGATCCGCCCATTGCCTGTTAAGCCCAAGACCA 1435
Db    1388 AATTCATCCAGCCTAGGGACACAACCATTATCCGACCAGTGCCTGTTAAGTCCAAGGCCA 1447

Qy    1436 AGGTTAAGCAGAGGGTGAGGGTGAAGACCGTGCCTTCTCTCTTCCAGGCCAAGCTGGACA 1495
Db    1448 AGAGTAAACAGAAAGTGAGAATAAAGACTGTGAGTTCTGCCTCCCAGCCCAAGCTGGATA 1507

Qy    1496 CCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACCAATGGCCAGACACATGCCTGCAACTT 1555
Db    1508 CCTTCTTATGTCAG----------------TGAACGGTGACCAGAGTGATGTTTGCAATTA 1552

Qy    1556 GTAGGTCAAGGACTGTCCA--GGCAGGGTTTGTGGACAGAGCCCTACTTTCGGGACCAGC 1613
Db    1553 GTGGGCCAAGACCTGGCTAACCGGAAGTGTTTTTGGAAGATGGCTCCTCTTGGACC---- 1608

Qy    1614 CTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTGAGTGTGACACCCAGAGACTGCTCTTC 1673
Db    1609 ----------------------------AGTCCAAGAGAGATGCCCAGAAAACACACTTC 1640

Qy    1674 CTGCCCTCACCCTGCCCCACTCCTACGACTGGGAGCTGACATGACCAGCCCACTGATCCT 1733
Db    1641 CTGTGTTCACTGCGCCCTGCACCAC--ACTGGGAAGCCACATGACCAGTTTACTGTTCCG 1698
```

```
Qy   1734 GTCAGCAGGTCCTGC-TCTGTTGCCAGGCTCTTGTTTATAGCCATGATCAGATGTGGTCA 1792
          ||||||||| ||| | ||    |  |||| | ||| |||||||| |  | ||| |||||
Db   1699 ATCAGCAGGGCCTACTTCCAGTTGCAGGGTTTTGCTTATAGCTACAACCAGGTGTGGCTG 1758

Qy   1793 GACTCTTTCTGGGCCTGGAGACCACGGTCACTTGTTGACTGTCTCTGTGGACCAG----- 1847
          ||||| || ||     |  ||| || |||||| |      ||   |   |||     |
Db   1759 GACTCCTTTTGTTTTTATAGAACAGGGTCACATTGACTCTAAGTGGATGGGAGTGCTGGA 1818

Qy   1848 AGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTCTACCTGCCTTTGACTTGCTTC 1907
          |  | |  |    ||   | | |||  ||   | || || |||||||    |||  | |
Db   1819 GGATCCTATGCAGGCTGGAGGACCCTGCGCTTGAACTCCTGCCTGCCTCCAGCTTATTGC 1878

Qy   1908 TAGCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGATAGACATGGGATGTATGGAGAG 1967
          | | |      |||  | |  | |   || |  |      ||  |    ||| |  |
Db   1879 TTGAAATTATGGGGTGAGGTGGTGATAGGGAAAGGTTGGGGAAGTTTTCTGTGTAAAATA 1938

Qy   1968 GATGGAAGATTTTCCCGAAAAAAAAAAAAAAAAAAAA 2004
          |  |     ||| |   ||||||||||||||||||
Db   1939 AAAAGGGATCTTTTCTTCAAAAAAAAAAAAAAAAAAA 1975
```

RESULT 5
ABL39714
ID ABL39714 standard; cDNA; 1781 BP.
XX
AC ABL39714;
XX
DT 10-MAY-2002 (first entry)
XX
DE Human NS cDNA sequence SEQ ID NO:24.
XX
KW Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW infertility; cardiovascular disease; coagulation disease; hypertension;
KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200206315-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-IL00653.
XX
PR 18-JUL-2000; 2000IL-0137345.
PR 15-DEC-2000; 2000IL-0140354.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Mintz L, Freilich S, Bernstein J;
XX
DR WPI; 2002-155037/20.
DR P-PSDB; ABB06060.
XX
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS Claim 1; Page 88; 290pp; English.
XX
CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive.
XX
SQ Sequence 1781 BP; 402 A; 498 C; 467 G; 405 T; 9 other;

Query Match 44.5%; Score 892.8; DB 24; Length 1781;
Best Local Similarity 80.9%; Pred. No. 8.9e-232;
Matches 1211; Conservative 0; Mismatches 42; Indels 243; Gaps 5;

```
Qy    748 AAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTTGTTTTCCTCCAGAAGCAAG 807
          |||  ||     ||  | ||  |  ||  |  ||| |  |||  | |        . |||
Db    287 AAGCTCTTCTTTGATCTTGCCCAGGAGGAGGAGAATGTCTTGGATGCAGAATTCTTAAAG 346

Qy    808 TTGCAGACAGTCTACTCTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAG 867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 TTGCAGACAGTCTACTCTGAATTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCAGAAG 406

Qy    868 GACTTACAGAGTGCTGACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAG 927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 GACTTACAGAGTGCTGACAAGGAAATCATGAGCCTGAAAAAGAAGCTAACGATGCTGCAG 466

Qy    928 GAAACCTTGAACCTGCCACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGC 987
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 GAAACCTTGAACCTGCCACCAGTGGCCAGTGAGACTGTCGACCGCCTGGTTTTAGAGAGC 526

Qy    988 CCAGCCCCTGTGGAGGTGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGAT 1047
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    527 CCAGCCCCTGTGGAGGTGAATCTGAAGCTCCGCCGGCCATCCTTCCGTGATGATATTGAT 586

Qy   1048 CTCAATGCTACCTTTGATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGT 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    587 CTCAATGCTACCTTTGATGTGGATACTCCCCCAGCCCGGCCCTCCAGCTCCCAGCATGGT 646

Qy   1108 TACTACGAAAAACTTTGCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAG 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    647 TACTACGAAAAACTTTGCCTAGAGAAGTCACACTCCCCAATTCAGGATGTCCCCAAGAAG 706

Qy   1168 ATATGCAAAGGCCCCAGGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGA 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    707 ATATGCAAAGGCCCCAGGAAGGAGTCCCAGCTCTCACTGGGTGGCCAGAGCTGTGCAGGA 766

Qy   1228 GAGCCAGATGAGGAACTGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGC 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    767 GAGCCAGATGAGGAACTGGTTGGTGCCTTCCCTATTTTTGTCCGGAATGCCATCCTAGGC 826

Qy   1288 CAGAAACAGCCCAAAAGGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACA 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    827 CAGAAACAGCCCAAAAGGCCCAGGTCAGAGTCCTCTTGCAGCAAAGATGTGGTAAGGACA 886

Qy   1348 GGCTTCGATGGGCTCGGTGGCCGGACAAAATTCATCCAGCCT---------------------- 1389
          ||||||||||||||||||||||||||||||||||||||||||
Db    887 GGCTTCGATGGGCTCGGTGGCCGGACAAAATTCATCCAGCCTGTWTCCTTTTYCTGTGAT 946

Qy   1390 ---------------------------------------------------------------- 1389
Db    947 GTTGAGCTGGGCTGAACTGTGGNGCCCATACTATGCCTCACCTATACCCCCAGCCTGCCA 1006
```

```
Qy   1390 ------------------------------------------------------------ 1389

Db   1007 GTGCCAACAGCCTCAGGTGTGNAGTGCTGGGCCTCATTCTGTGTNTCAGCGCCGCCCACC 1066

Qy   1390 ------------------------------------------------------------ 1389

Db   1067 CCCGGCTCTGAGCCTTTCTTGCCCTGNATNGNAGTCCTTTCTCGCCCTGCATGGCAGTCC 1126

Qy   1390 ----------------------------------------ACTGACACAGTCATGATCCG 1409
                                                    ||||||||||||||||||||
Db   1127 TTTCTCCATGCCTCTTCCTTAACNGCTCCTCCAAACTAAAACTGACACAGTCATGATCCG 1186

Qy   1410 CCCATTGCCTGTTAAGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACCGTGCC 1469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1187 CCCATTGCCTGTTAAGCCCAAGACCAAGGTTAAGCAGAGGGTGAGGGTGAAGACCGTGCC 1246

Qy   1470 TTCTCTCTTCCAGGCCAAGCTGGACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACC 1529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247 TTCTCTCTTCCAGGCCAAGCTGGACACCTTCCTGTGGTCGTGAGAACAGTGAGTCTGACC 1306

Qy   1530 AATGGCCAGACACATGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGG--TTTGTG 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db   1307 AATGGCCAGACACATGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCAGGGGTTTTGTG 1366

Qy   1588 GACAGAGCCCTACTTTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTG 1647
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1367 GACAGAGCCCCACTTTCGGGACCAGCCTGAGGTGTAAGGGCAGACAAACAGGTGAGGGTG 1426

Qy   1648 AGTGTGACACCCAGAGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGA 1707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1427 AGTGTGACACCCAGAGACTGCTCTTCCTGCCCTCACCCTGCCCCACTCCTACGACTGGGA 1486

Qy   1708 GCTGACATGACCAGCCCACTGATCCTGTCAGCAGGTCCTGCT-CTGTTGCCAGGCTCTTG 1766
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||
Db   1487 GCTGACATGACCAGCCCACTGATCCTGTCAGCAGGTCCTGCTCCTGTTGCCAGGCTCCTG 1546

Qy   1767 TTTATAGCCATGATCAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTG 1826
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 TTTATAGCCATGATCAGATGTGGTCAGACTCTTTCTGGGCCTGGAGACCACGGTCACTTG 1606

Qy   1827 TTGACTGTCTCTGTGGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTC 1886
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 TTGACTGTCTCTGTGGACCAGAGTGCTTGAGGCATCTCAGGCAGCCTCAGCCCAAGCTTC 1666

Qy   1887 TACCTGCCTTTGACTTGCTTCTA-GCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGA 1945
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1667 TACCTGCCTTTGACTTGCTTCTAGGCATAGCCTGGGCCAAGCAGGGTGGGGAATGGAGGA 1726

Qy   1946 TAGACATGGGATGTATGGAGAGGATGGAAGATTTTCCCGAAAAAAAAAAAAAAAAA 2001
          ||| ||||||||||||||||||||||||||||||  | |||| |||| |||||
Db   1727 TAG-CATGGGATGTATGGAGAGGATGGAAGATTTTCATGTAAAATAAAATTAAAAA 1781


RESULT 6
ABA95683
ID   ABA95683 standard; DNA; 7542 BP.
XX
AC   ABA95683;
XX
DT   03-APR-2002  (first entry)
XX
DE   Human protein kinase gene.
XX
KW   Human; protein kinase; enzyme; gene; brain; lung; hippocampus;
KW   calmodulin-binding kinase; gene therapy; chromosome 3;
KW   single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(234,A)
FT                   /*tag= a
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(328,C)
FT                   /*tag= b
FT                   /standard_name= "Single nucleotide polymorphism"
FT   CDS             1473..4544
FT                   /*tag= c
FT                   /product= "Human protein kinase"
FT                   /transl_except= (pos:4275..4277, aa:Val)
FT                   /note= "contains 9 introns"
FT   exon            1473..1567
FT                   /*tag= d
FT                   /number= 1
FT   variation       replace(1499,T)
FT                   /*tag= e
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          1568..1684
FT                   /*tag= f
FT                   /number= 1
FT   exon            1685..1816
FT                   /*tag= g
FT                   /number= 2
FT   intron          1817..1996
FT                   /*tag= h
FT                   /number= 2
FT   exon            1997..2071
FT                   /*tag= i
FT                   /number= 3
FT   intron          2072..2284
FT                   /*tag= j
FT                   /number= 3
FT   exon            2285..2423
FT                   /*tag= k
FT                   /number= 4
FT   intron          2424..2561
FT                   /*tag= l
FT                   /number= 4
FT   exon            2562..2682
FT                   /*tag= m
FT                   /number= 5
FT   intron          2683..2843
FT                   /*tag= n
FT                   /number= 5
FT   exon            2844..2919
FT                   /*tag= o
FT                   /number= 6
FT   intron          2920..3009
FT                   /*tag= p
FT                   /number= 6
FT   exon            3010..3146
FT                   /*tag= q
FT                   /number= 7
FT   variation       replace(3022,T)
FT                   /*tag= r
FT                   /standard_name= "Single nucleotide polymorphism"
FT   intron          3147..3312
FT                   /*tag= s
FT                   /number= 7
FT   variation       replace(3295,T)
FT                   /*tag= t
FT                   /standard_name= "Single nucleotide polymorphism"
FT   exon            3313..3391
FT                   /*tag= u
FT                   /number= 8
FT   intron          3392..3603
FT                   /*tag= v
FT                   /number= 8
FT   exon            3604..3691
FT                   /*tag= w
FT                   /number= 9
FT   intron          3692..3980
FT                   /*tag= x
FT                   /number= 9
FT   exon            3981..4541
FT                   /*tag= y
```

```
FT                   /number= 10
FT   variation       replace(4677,T)
FT                   /*tag= z
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(5583,Y)
FT                   /*tag= aa
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(6242,C)
FT                   /*tag= ab
FT                   /standard_name= "Single nucleotide polymorphism"
FT   variation       replace(7264,G)
FT                   /*tag= ac
FT                   /standard_name= "Single nucleotide polymorphism"
XX
PN   WO200192492-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17327.
XX
PR   30-MAY-2000; 2000US-207281P.
PR   12-DEC-2000; 2000US-0734030.
XX
PA   (APPL-) APPLERA CORP.
XX
PI   Yan C,  Wei M,  Ketchum K,  Merkulov G,  Beasley EM;
XX
DR   WPI; 2002-097770/13.
DR   P-PSDB; AAM48279.
XX
PT   New calmodulin-binding kinase peptides and nucleic acid encoding the
PT   peptides, useful as models for developing human therapeutic targets or
PT   in screening for compounds that modulate kinase  -
XX
PS   Claim 4; Fig 3; 75pp; English.
XX
CC   The present sequence is a human protein kinase gene. The protein kinase
CC   coding sequence (see ABA95682) is expressed in the brain (both infant
CC   and adult brain), lung and hippocampus. The protein kinase is related to
CC   the calmodulin-binding kinase subfamily. The protein kinase and its
CC   coding sequence can be used as models for the development of human
CC   therapeutic targets, in the identification of therapeutic proteins, and
CC   serve as targets for the development of human therapeutic agents that
CC   modulate kinase activity in cells and tissues that express the kinase. In
CC   addition, the protein kinase coding sequence can be used for treating a
CC   disorder associated with nucleic acid expression of the kinase gene,
CC   particularly biological and pathological processes that are mediated by
CC   the kinase in cells and tissues that express it, as antisense constructs
CC   to control kinase gene expression in cells, tissues or organisms, and in
CC   gene therapy. The protein kinase gene maps to chromosome 3.
XX
SQ   Sequence 7542 BP; 1612 A; 1977 C; 2156 G; 1797 T; 0 other;

  Query Match             8.8%;  Score 176.4;  DB 24;  Length 7542;
  Best Local Similarity   95.1%;  Pred. No. 5.3e-37;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels    4;  Gaps    1;

Qy       1 GTGCGGTGGAGCGAAATTTGAAGCAAGCGGAGGCGGGGCG----CTCTACGAAGCCGGAC 56
           |||| | ||||| |||||||||||||||||||||| || |     ||||||||||||||
Db    7302 GTGCTGCGGAGCCAAATTTGAAGCAAGCGGAGGCGCGGGGGCGCGTCTACGAAGCCGGAC 7361

Qy      57 CTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCCTATCCGTGC 116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7362 CTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCCTATCCGTGC 7421

Qy     117 TCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGCCATCCACTG 176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7422 TCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGCCATCCACTG 7481

Qy     177 CGGCCACACCTTCCACTTGCAGTG 200
           |||||||||||||||||||||||
Db    7482 CGGCCACACCTTCCACTTGCAGTG 7505
```

```
RESULT 7
AAS27719/c
ID   AAS27719 standard; DNA; 148 BP.
XX
AC   AAS27719;
XX
DT   07-NOV-2001  (first entry)
XX
DE   DNA encoding novel signal transduction pathway protein, Seq ID 1379.
XX
KW   Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW   antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW   immune system disorder; rheumatoid arthritis; inflammatory condition;
KW   organ transplant rejection; infection; hepatitis C; blood disorder;
KW   sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW   cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW   reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW   acquired immune deficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200154733-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01312.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
```

PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 1379; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX
SQ Sequence 148 BP; 38 A; 33 C; 56 G; 21 T; 0 other;

Query Match 7.4%; Score 148; DB 22; Length 148;
Best Local Similarity 100.0%; Pred. No. 4e-30;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 AAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCC 107
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 148 AAGCCGGACCTGTAGCAGTTTCTTTGGCTGCCTGGGCCCCTTGAGTCCAGCCATCATGCC 89

Qy 108 TATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGC 167
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 88 TATCCGTGCTCTGTGCACTATCTGCTCCGACTTCTTCGATCACTCCCGCGACGTGGCCGC 29

Qy 168 CATCCACTGCGGCCACACCTTCCACTTG 195
|||||||||||||||||||||||||||||
Db 28 CATCCACTGCGGCCACACCTTCCACTTG 1

```
RESULT 8
ABN41009
ID   ABN41009 standard; DNA; 60 BP.
XX
AC   ABN41009;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:13757.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 13757; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 16 A; 16 C; 16 G; 12 T; 0 other;

  Query Match             3.0%;  Score 60;  DB 24;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2e-06;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1519 TGAGTCTGACCAATGGCCAGACACATGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCA 1578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGAGTCTGACCAATGGCCAGACACATGCCTGCAACTTGTAGGTCAAGGACTGTCCAGGCA 60
```

```
RESULT 9
AAA30290
ID   AAA30290 standard; DNA; 3489 BP.
XX
AC   AAA30290;
XX
DT   11-SEP-2000  (first entry)
XX
DE   Kaposi's sarcoma-associated herpesvirus LANA gene.
XX
KW   Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
KW   latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
KW   Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
KW   Kaposi's sarcoma; primary effusion lymphoma; PEL;
KW   human immunodeficiency virus; HIV; multicentric Castleman's disease; ds.
XX
OS   Kaposi's sarcoma-associated herpesvirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3489
FT                   /*tag= a
FT                   /product= "LANA"
FT   misc_signal     40..50
FT                   /*tag= b
FT                   /note= "nuclear localisation signal, NLS"
FT   misc_signal     190..210
FT                   /*tag= c
FT                   /note= "nuclear localisation signal, NLS"
XX
PN   WO200029626-A1.
XX
PD   25-MAY-2000.
XX
PF   19-NOV-1999;   99WO-US27508.
XX
PR   19-NOV-1998;   98US-0109422.
PR   21-APR-1999;   99US-0298568.
XX
PA   (KIEF/) KIEFF E D.
PA   (BALL/) BALLESTAS M E.
PA   (KAYE/) KAYE K M.
XX
PI   Kieff ED,  Ballestas ME,  Kaye KM;
XX
DR   WPI; 2000-387829/33.
DR   P-PSDB; AAY96255.
XX
PT   Treating or preventing a disease associated with rhodino virus
PT   infection in a mammal which includes Kaposi's Sarcoma and Primary
PT   Effusion Lymphoma -
XX
PS   Disclosure; Fig 6; 70pp; English.
XX
CC   The present sequence is the Kaposi's sarcoma-associated herpesvirus,
CC   (KSHV) latency-associated nuclear antigen (LANA) gene. KSHV is also known
CC   as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or
CC   gamma-2 herpes virus class. The LANA protein is necessary for the
CC   efficient persistence of rhadino virus DNA in mammalian cells. Persistent
CC   rhadino virus infection is implicated in a variety of diseases e.g.
CC   Kaposi's Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
CC   Castleman's disease. In addition, KS is a common malignancy in HIV
CC   patients. KSHV persists in host cells in a latent form. One of the few
CC   genes expressed from the latent viral DNA is LANA. LANA associates with
CC   both human chromosomes and with the rhadino virus cis-acting element
CC   (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
CC   "tied" to the host chromosomes. This allows the viral DNA to persist in
CC   the host cell. The present sequence may be used to screen and identify
CC   molecules that inhibit LANA interaction with RVCAE, thereby interfering
CC   with the latency cycle of this virus. Potential antiviral treatments for
CC   the above mentioned diseases may therefore be based on LANA deregulation.
XX
```

```
SQ   Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

  Query Match             2.6%;  Score 52;  DB 21;  Length 3489;
  Best Local Similarity   48.9%;  Pred. No. 0.0021;
  Matches  139;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     505 CTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAG 564
           | | |  |||||  | || |    |||||||||||| | ||     | |    || |||
Db    2212 CAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAG 2271

Qy     565 GAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAG 624
           |||  ||  | ||    |    ||  | ||    | ||| ||   ||||       |  ||
Db    2272 GAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG 2331

Qy     625 AGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG 684
             |||   |  ||| | ||||||  |   |  ||  | |  | |    ||   | | |
Db    2332 GATCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA 2391

Qy     685 GAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCA 744
           ||  ||| ||    | | |  |   || | ||||  ||| | |   |  |||
Db    2392 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAG 2451

Qy     745 CGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTT 788
             |||     | | ||| |   ||  ||| |  | |||| ||
Db    2452 TTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTT 2495


RESULT 10
AAF82901
ID   AAF82901 standard; DNA; 3489 BP.
XX
AC   AAF82901;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW   Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW   Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW   KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS   Kaposi's sarcoma associated herpesvirus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3489
FT                   /*tag=  a
XX
PN   WO200125484-A2.
XX
PD   12-APR-2001.
XX
PF   29-SEP-2000; 2000WO-US26908.
XX
PR   01-OCT-1999;   99US-0410399.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Robertson ES,  Cotter MA;
XX
DR   WPI; 2001-281736/29.
DR   P-PSDB; AAB62331.
XX
PT   A composition for use in gene therapy comprises an expression vector
PT   that  includes a nucleic acid sequence encoding a nucleic acid binding
PT   protein  -
XX
PS   Disclosure; Fig 9A; 60pp; English.
XX
CC   The invention provides a composition comprising nucleic acid, histone H1
CC   protein and expression vector operationally encoding a protein suitable
CC   for tethering the nucleic acid to the histone H1 protein, where the
CC   tethering protein is LANA. The composition is useful in aiding the
CC   retention of the viral DNA in the host cell. The viral vector encodes a
CC   protein suitable for tethering DNA to Histone H1. Methods for screening
CC   for compounds which are agonistic or antagonistic for the tethering of
CC   viral proteins to histone H1 and DNA binding sites are useful for
CC   developing the method of viral transfer. The composition has applications
CC   to gene therapy, including the treatment of multiple sclerosis,
CC   Parkinson's disease, Huntington disease and diabetes. The present
CC   sequence represents the nucleotide sequence of the Kaposi's sarcoma
CC   associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC   antigen (LANA), which acts as a tethering protein.
XX
SQ   Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

  Query Match             2.6%;  Score 52;  DB 22;  Length 3489;
  Best Local Similarity   48.9%;  Pred. No. 0.0021;
  Matches  139;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     505 CTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAG 564
           | | |  |||||  | || |    |||||||||||| | ||     | |    || |||
Db    2212 CAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAG 2271

Qy     565 GAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAG 624
           |||  ||  | ||    |    ||  | ||    | ||| ||   ||||       |  ||
Db    2272 GAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG 2331

Qy     625 AGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG 684
             |||   |  ||| | ||||||  |   |  ||  | |  | |    ||   | | |
Db    2332 GATCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA 2391

Qy     685 GAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCA 744
           ||  ||| ||    | | |  | |   |  || | ||||  ||| | |   |  |||
Db    2392 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAG 2451

Qy     745 CGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTT 788
             |||     | | ||| |   ||  ||| |  | |||| ||
Db    2452 TTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTT 2495


RESULT 11
ABA93487
ID   ABA93487 standard; DNA; 3489 BP.
XX
AC   ABA93487;
XX
DT   25-APR-2002  (first entry)
XX
DE   Kaposi's sarcoma-associated herpesvirus LANA protein encoding DNA.
XX
KW   Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;
KW   KSHV terminal repeat; rhadino virus cis acting element; episome;
KW   primary effusion lymphoma; latency-associated nuclear antigen;
KW   gene therapy; gene transfer; gene; ds.
XX
OS   Human herpesvirus 8.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3489
FT                   /*tag=  a
FT                   /product= "LANA protein"
FT                   /note= "latency-associated nuclear antigen"
XX
PN   US6322792-B1.
XX
PD   27-NOV-2001.
XX
PF   21-APR-1999;   99US-0298568.
XX
PR   19-NOV-1998;   98US-109422P.
XX
PA   (KIEF/) KIEFF E D.
XX
PI   Kieff ED,  Ballestas ME,  Kaye KM;
XX
DR   WPI; 2002-153769/20.
```

```
DR   P-PSDB; ABB05621.
XX
PT   System for episomal retention of plasmids in mammalian cells, useful in
PT   gene therapy, comprises rhadinoviral LANA and RVCAE sequences -
XX
PS   Claim 1; Fig 6; 27pp; English.
XX
CC   The present invention describes a system (A) for maintaining a plasmid
CC   as an episome in mammalian cells, comprising the rhadinoviral sequence
CC   LANA (latency-associated nuclear antigen) of 3489 base pairs (see
CC   ABA93487, S1) expressed in the cell, and the rhadinoviral sequence RVCAE
CC   (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2)
CC   present in the plasmid. Also describes is a method for maintaining a
CC   closed circular DNA in a cell by expressing (S1) in the cells and having
CC   (S2) as a cis-acting and maintenance sequence in the DNA. (A) is
CC   particularly used in gene therapy (or other gene transfer applications)
CC   that uses mammalian cells in which LANA is expressed. (A) improves
CC   persistence of gene therapy vectors in cells. The present sequence
CC   encodes Kaposi's sarcoma-associated herpesvirus (KSHV, also called human
CC   herpesvirus 8) LANA protein, which is used in the exemplification of the
CC   present invention.
XX
SQ   Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

  Query Match             2.6%;  Score 52;  DB 24;  Length 3489;
  Best Local Similarity   48.9%;  Pred. No. 0.0021;
  Matches  139;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     505 CTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAG 564
           | | |  |||||  | || |    ||||||||||||| | ||     |  |    || |||
Db    2212 CAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAG 2271

Qy     565 GAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAG 624
           |||  ||  | ||    |    ||  | ||    | ||| ||   ||||      |  ||
Db    2272 GAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG 2331

Qy     625 AGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG 684
              |||   |  ||| | ||||||  |   |  ||  | |  | |    ||   | | |
Db    2332 GATCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA 2391

Qy     685 GAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCA 744
           ||  ||| ||    | | |  | |   |   || | ||||  ||| | |   |  |||
Db    2392 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAG 2451

Qy     745 CGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTT 788
              |||     | | ||| |   ||  ||| |  | |||| ||
Db    2452 TTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTT 2495


RESULT 12
AAV73805/c
ID   AAV73805 standard; DNA; 32207 BP.
XX
AC   AAV73805;
XX
DT   25-FEB-1999  (first entry)
XX
DE   KSHV LUR DNA (nucleotides 105,301-137,507).
XX
KW   Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
KW   dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;
KW   diagnosis; treatment; HHV8; capsid protein IV; tegument protein IV;
KW   glycoprotein; kaposin; cyclin D; immediate early protein; IEP; OX-2;
KW   v-adh; G-protein coupled receptor; FGARAT; ds.
XX
OS   Kaposi's sarcoma-associated herpesvirus.
XX
PN   US5849564-A.
XX
PD   15-DEC-1998.
XX
PF   29-NOV-1996;  96US-0770379.
XX
PR   29-NOV-1996;  96US-0770379.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Bohenzky RA, Chang Y, Edelman IS, Moore PS, Russo JJ;
XX
DR   WPI; 1999-069741/06.
XX
PT   Kaposi's sarcoma-associated herpes virus nucleic acid - encodes
PT   di:hydro:folate reductase and is useful for treatment, prophylaxis
PT   or diagnosis of Kaposi's sarcoma
XX
PS   Disclosure; Column 155-182; 109pp; English.
XX
CC   This sequence is a fragment of the Kaposi's sarcoma-associated
CC   herpesvirus (KSHV) LUR (long unique region). This fragment contains
CC   coding regions for ORF65 which encodes capsid protein IV, ORF66, ORF67
CC   which encodes tegument protein IV, ORF68 which encodes a glycoprotein,
CC   ORF69, K12 which encodes kaposin, K13, ORF72 which encodes cyclin D,
CC   ORF73 which encodes immediate early protein (IEP), K14 which encodes
CC   OX-2 (v-adh), ORF74 which encodes G-protein coupled receptor, ORF75
CC   which encodes tegument protein/FGARAT, K15. KSHV is a new human
CC   Herpesvirus (HHV8) believed to cause Kaposi's sarcoma (KS) which is the
CC   most common form of neoplasm occurring in persons with acquired immune
CC   deficiency syndrome (AIDS). The DHFR protein is useful for vaccination,
CC   prophylaxis, diagnosis and treatment of a subject with Kaposi's sarcoma
CC   and for detecting expression of a DNA virus associated with Kaposi's
CC   sarcoma in a cell.
XX
SQ   Sequence 32207 BP; 7229 A; 9156 C; 8713 G; 7109 T; 0 other;

  Query Match             2.6%;  Score 52;  DB 20;  Length 32207;
  Best Local Similarity   48.9%;  Pred. No. 0.0064;
  Matches  139;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     505 CTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAG 564
           | | |  |||||  | || |    ||||||||||||| | ||     |  |    || |||
Db   19785 CAGGATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAG 19726

Qy     565 GAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAG 624
           |||  ||  | ||    |    ||  | ||    | ||| ||   ||||      |  ||
Db   19725 GAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG 19666

Qy     625 AGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG 684
              |||   |  ||| | ||||||  |   |  ||  | |  | |    ||   | | |
Db   19665 GATCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA 19606

Qy     685 GAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCA 744
           ||  ||| ||    | | |  | |   |   || | ||||  ||| | |   |  |||
Db   19605 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAG 19546

Qy     745 CGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTT 788
              |||     | | ||| |   ||  ||| |  | |||| ||
Db   19545 TTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTT 19502


RESULT 13
AAV19941/c
ID   AAV19941 standard; DNA; 137507 BP.
XX
AC   AAV19941;
XX
DT   03-AUG-1998  (first entry)
XX
DE   KSHV long unique coding region and terminal repeat.
XX
KW   KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;
KW   interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;
KW   complement-binding protein; glycoprotein; capsid protein IV; infection;
KW   immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;
KW   lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;
KW   HIV immune status; anti-inflammatory agent; therapy; ds.
XX
```

```
OS   Kaposi's sarcoma-associated herpes virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1142..2794
FT                   /*tag= a
FT                   /product= complement-binding protein
FT   CDS             8699..11236
FT                   /*tag= b
FT                   /product= glycoprotein B
FT   CDS             complement (17261..17875)
FT                   /*tag= c
FT                   /product= interleukin 6
FT   CDS             complement (21548..21832)
FT                   /*tag= d
FT                   /product= macrophage inflammatory protein II
FT   CDS             complement (27137..27424)
FT                   /*tag= e
FT                   /product= interferon regulatory factor 1
FT   CDS             28661..29741
FT                   /*tag= f
FT                   /product= protein T1.1
FT   CDS             complement (58976..60175)
FT                   /*tag= g
FT                   /product= glycoprotein M
FT   CDS             complement (69412..69915)
FT                   /*tag= h
FT                   /product= glycoprotein L
FT   CDS             complement (88410..88910)
FT                   /*tag= i
FT                   /product= interferon regulatory factor 2
FT   CDS             89600..90541
FT                   /*tag= j
FT                   /product= interferon regulatory factor 3
FT   CDS             90173..90643
FT                   /*tag= k
FT                   /product= glycoprotein X
FT   CDS             complement (93636..94127)
FT                   /*tag= l
FT                   /product= interferon regulatory factor 4
FT   CDS             complement (111931..112443)
FT                   /*tag= m
FT                   /product= capsid protein IV
FT   CDS             complement (123808..127296)
FT                   /*tag= n
FT                   /product= immediate early protein
XX
PN   WO9804576-A1.
XX
PD   05-FEB-1998.
XX
PF   22-JUL-1997;    97WO-US13346.
XX
PR   29-NOV-1996;    96US-0757669.
PR   25-JUL-1996;    96US-0686243.
PR   25-JUL-1996;    96US-0686349.
PR   25-JUL-1996;    96US-0686350.
PR   25-JUL-1996;    96US-0687253.
PR   25-JUL-1996;    96US-0688814.
PR   05-SEP-1996;    96US-0708678.
PR   10-OCT-1996;    96US-0728323.
PR   13-NOV-1996;    96US-0747887.
PR   13-NOV-1996;    96US-0748640.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Bohenzky RA,  Chang Y,  Edelman IS,  Moore PS,  Russo JJ;
XX
DR   WPI; 1998-130615/12.
XX
PT   New nucleic acid encoding Kaposi's sarcoma associated herpes virus
PT   proteins - useful for, e.g. detecting levels of HHV8 in, and
PT   preparation of vaccines for treatment of, HIV patients
XX
PS   Example 2; Page 135-203; 230pp; English.
XX
CC   This sequence represents the long unique region and terminal repeat of
CC   the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
CC   as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
CC   invention which encode KSHV polypeptides selected from: (a) viral
CC   macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
CC   (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
CC   (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
CC   encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
CC   by it, and antibodies (Ab) specific for the proteins are useful for
CC   detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
CC   fluids or tissue samples. HHV8 infections can be treated with antisense
CC   or triplex forming molecules or agents that bind specifically to the
CC   protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
CC   while the protein can be used in protective vaccines. Ab may also be used
CC   to differentiate between lymphomas, and HHV8 may be implicated in many
CC   other lymphoproliferative diseases such as lymphomas, leukaemia,
CC   splenomegaly and mycosis fungoides. Cells and animals containing the
CC   nucleic acid are useful for drug screening. HHV8-derived peptides can be
CC   used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
CC   can be inhibited with methotrexate. These can also be used to determine
CC   the immune status of a patient infected with HIV. HHV8 derived protein
CC   viral MIP III may be used as an anti-inflammatory agent for,
CC   e.g. treating rheumatoid arthritis. This sequence is stated as containing
CC   81 open reading frames.
XX
SQ   Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 other;

  Query Match             2.6%;  Score 52;  DB 19;  Length 137507;
  Best Local Similarity   48.9%;  Pred. No. 0.013;
  Matches  139;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy     505 CTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGGATGAGACCAAACAAGCACAAGAG 564
           | | |  ||||| | || |    ||||||||||||| | ||     |  |    || |||
Db  125085 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAACAGGAGCAGCAGGAG 125026

Qy     565 GAGGCGGGCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTTCTACTCCAG 624
           |||  ||  | ||    |    ||  | ||    | ||| ||   ||||      |  ||
Db  125025 GAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG 124966

Qy     625 AGCCAGCTCCCTGAGGTGGAGGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG 684
              |||   |  ||| | ||||||  |   |  ||  | |  | |    ||   | | |
Db  124965 GATCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTA 124906

Qy     685 GAACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTAAAAGAGGCA 744
           ||  ||| ||    | | |  | |   |   || | ||||  ||| | |   |  |||
Db  124905 GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAG 124846

Qy     745 CGGAAGGCCTCAGGGGAGGTGGCTGACAAGCTGAGGAAGGATTT 788
              |||     | | ||| |   ||  ||| |  | |||| ||
Db  124845 TTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTT 124802


RESULT 14
AAS60947
ID   AAS60947 standard; cDNA; 4246 BP.
XX
AC   AAS60947;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human cancer agent-resistance marker #606.
XX
KW   Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW   squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW   lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;
KW   Hodgkin's disease; glioma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200179556-A2.
XX
```

```
PD   25-OCT-2001.
XX
PF   13-APR-2001; 2001WO-US12132.
XX
PR   14-APR-2000; 2000US-197538P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lillie J,  Brown JL,  Bolt A,  Van Huffel C;
XX
DR   WPI; 2001-602933/68.
XX
PT   Novel nucleic acid, used as a marker to determine the effectiveness of
PT   using TAXOL to treat cancer cell growth in individuals -
XX
PS   Claim 1; Page 450-451; 527pp; English.
XX
CC   The invention relates to 1046 novel nucleic acids which are used as
CC   markers for determining the sensitivity of a cancer cell to the
CC   anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
CC   they are shown to express one of the 242 sensitivity markers or the
CC   cells are shown not to express one of the 804 resistance markers.
CC   The methods can be used to determine the effectiveness of TAXOL
CC   in the treatment of cancer cell growth in an individual. The markers
CC   can be used as targets in developing anti-cancer agents such as
CC   chemotherapeutic compounds. The markers can also be used as targets in
CC   developing treatments for cancer, particularly those cancers which
CC   display resistance to agents and exhibit expression of the markers. The
CC   anticancer agents developed by the novel method can be used to treat
CC   cancer. Probes based on the markers can be used to detect transcripts or
CC   genomic sequences corresponding to the markers, in the identification of
CC   cells or tissues which mis-express the protein. Cancers which may
CC   be targeted include carcinoma (e.g. squamous cell carcinoma),
CC   sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
CC   lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's disease and
CC   tumours (e.g. glioma). The present sequence is one of the 1046
CC   novel cancer cell markers.
XX
SQ   Sequence 4246 BP; 1217 A; 1019 C; 1183 G; 827 T; 0 other;

  Query Match             2.6%;  Score 51.8;  DB 22;  Length 4246;
  Best Local Similarity   47.9%;  Pred. No. 0.0027;
  Matches  149;  Conservative    0;  Mismatches  162;  Indels    0;  Gaps    0;

Qy     297 TCTTGCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGAATGAACTGGACAA 356
           ||  |  ||| ||||||| | |    ||   |  |   | |  |||| || ||
Db     765 TCAGGAGCAGCAGGAGGAGGAGGAGGAGGCCAGGAAGAACCTGGAGAAGCAAGTGCTGGC 824

Qy     357 TGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCCAGGTCATCATCGA 416
             |     ||||| |  |  | |   | ||| |   |   |||  || ||  |  || ||
Db     825 CCTGCAGTCCCAGTTGGCTGATACCAAGAAGAAAGTAGATGACGACCTGGGAACAATTGA 884

Qy     417 CACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTCTGCAGCAGGCCTT 476
           | ||||   ||  |   | |  | |           | ||   | |||   |||  | |
Db     885 AAGTCTGGAAGAAGCCAAGAAGAAGCTTCTGAAGGACGCGGAGGCCCTGAGCCAGCGCCT 944

Qy     477 GGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCA 536
           ||   ||   |   ||  ||    || |||| | |||   | |||   | |  ||||| |
Db     945 GGAGGAGAAGGCACTGGCGTATGACAAACTGGAGAAGACCAAGAACCGCCTGCAGCAGGA 1004

Qy     537 GCAGGATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGAGCAAGATGAAGAC 596
           || ||| ||   ||     |  |  ||  |   | ||| | ||     |||  || |||
Db    1005 GCTGGACGACCTCACGGTGGACCTGGACCACCAGCGCCAGGTCGCCTCCAACTTGGAGAA 1064

Qy     597 CATGGAGCAGA 607
           | | || |||
Db    1065 GAAGCAGAAGA 1075


RESULT 15
ABL66905
ID   ABL66905 standard; DNA; 7596 BP.
XX
AC   ABL66905;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:5242.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   27-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   28-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a
PT   chemical agent to be tested for anti-neoplastic activity, and
PT   determining a change in expression of a gene of a signature gene set  -
XX
```

PS Claim 1; SEQ ID 5242; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 7596 BP; 2283 A; 1697 C; 2021 G; 1595 T; 0 other;

Query Match 2.6%; Score 51.8; DB 24; Length 7596;
Best Local Similarity 47.9%; Pred. No. 0.0036;
Matches 149; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
Qy    297 TCTTGCCCAGGAGGAGGAGAATGTCTTGGATCGAGAATTCTTAAAGAATGAACTGGACAA 356
          ||  |  ||| ||||||||| | |    ||   |  |   | |  ||||  || ||
Db   4123 TCAGGAGCAGCAGGAGGAGGAGGAGGCCAGGAAGAACCTGGAGAAGCAAGTGCTGGC 4182

Qy    357 TGTCAGAGCCCAGCTTTCCCAGAAAGACAAGGAGAAACGAGACAGCCAGGTCATCATCGA 416
            |     ||||| |  |  | |   | ||| |   |   |||  || ||  |  || ||
Db   4183 CCTGCAGTCCCAGTTGGCTGATACCAAGAAGAAAGTAGATGACGACCTGGGAACAATTGA 4242

Qy    417 CACTCTGCGGGATACGCTGGAAGAACGCAATGCTACTGTGGTATCTCTGCAGCAGGCCTT 476
           | ||||   ||  |   | |  | |           | ||   | |||   |||  | |
Db   4243 AAGTCTGGAAGAAGCCAAGAAGAAGCTTCTGAAGGACGCGGAGGCCCTGAGCCAGCGCCT 4302

Qy    477 GGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCA 536
          ||   ||   |   ||  ||    || |||| | |||   | |||   | |  ||||| |
Db   4303 GGAGGAGAAGGCACTGGCGTATGACAAACTGGAGAAGACCAAGAACCGCCTGCAGCAGGA 4362

Qy    537 GCAGGATGAGACCAAACAAGCACAAGAGGAGGCGGGCCGGCTCAGGAGCAAGATGAAGAC 596
          || ||| ||   ||     |  |  ||  |   | ||| | ||     |||  || |||
Db   4363 GCTGGACGACCTCACGGTGGACCTGGACCACCAGCGCCAGGTCGCCTCCAACTTGGAGAA 4422

Qy    597 CATGGAGCAGA 607
           | | || |||
Db   4423 GAAGCAGAAGA 4433
```

Search completed: December 13, 2002, 03:05:42
Job time : 432 secs